Early Journal Content on JSTOR, Free to Anyone in the World

This article is one of nearly 500,000 scholarly works digitized and made freely available to everyone in the world by JSTOR.

Known as the Early Journal Content, this set of works include research articles, news, letters, and other writings published in more than 200 of the oldest leading academic journals. The works date from the mid-seventeenth to the early twentieth centuries.

We encourage people to read and share the Early Journal Content openly and to tell others that this resource exists. People may post this content online or redistribute in any way for non-commercial purposes.

Read more about Early Journal Content at http://about.jstor.org/participate-jstor/individuals/early-journal-content.

JSTOR is a digital library of academic journals, books, and primary source objects. JSTOR helps people discover, use, and build upon a wide range of content through a powerful research and teaching platform, and preserves this content for future generations. JSTOR is part of ITHAKA, a not-for-profit organization that also includes Ithaka S+R and Portico. For more information about JSTOR, please contact support@jstor.org.

THE
AMERICAN NATURALIST

VOL. LIV. March–April, 1920 No. 631

ARE THE FACTORS OF HEREDITY ARRANGED IN A LINE?

DR. H. J. MULLER
COLUMBIA UNIVERSITY

IN the February (1919) number of the *Proceedings of the National Academy of Sciences,* Professor Castle states that he has "shown that the arrangement of the genes in the sex-chromosome of *Drosophila ampelophila* is probably not linear, and a method has been developed for constructing a model of the experimentally determined linkage relationships."[1] This declaration is so widely at variance with the conclusions jointly agreed upon by all *Drosophila* workers, that the arguments or assumptions which it involves would seem to call for careful examination. It may be stated at the outset that the principle upon which Professor Castle constructs his models appears exceedingly direct and simple—it is merely to make a figure such that the distances between all the points represented on it are exactly proportional to the frequencies of separation actually found between the respective factors in the most reliable experiments. If this is done, Castle contends, the models will be three-dimensional instead of linear in shape.

1. The first argument which Castle gives against the view that the groups of genes (which he admits, at least

---

[1] Sturtevant, Bridges and Morgan also have published a defense of the view of linear linkage, in the *Proceedings* of the National Academy of Sciences (5, 1919, pp. 168–173) and Professor Castle has just replied to them in the same journal (5, 1919, pp. 501–506). It is believed that the present paper, although written and accepted for publication in the NATURALIST, previously to this article, meets all the points therein brought forward.

for purposes of argument, to be in the chromosomes) are linear, is that "it is doubtful . . . whether an elaborate organic molecule ever has a simple string-like form." This argument is therefore based upon the unique assumption that the whole chromosome (or that part of it containing the genes) consists of one huge molecule. Later, he speaks still more explicitly of this "chromosome molecule" and says, "the duplex linkage systems of a germ cell at the reduction division must be . . . twin organic molecules," so that "a purely mechanical theory [of crossing over] seems inadequate to account for interchange of equivalent parts between them." The argument may therefore be paraphrased as follows: since (1) the whole group of genes is but a single organic molecule, and since (2) an organic molecule can not be linear, then it must follow that (3) the group of genes is not linear, and that the theory of crossing over is therefore erroneous. Although the premises of this argument are both entirely gratuitous, it must be admitted that there is no flaw in the reasoning, once the premises are admitted.

2. The second argument brought forward against the linear arrangement of genes is that, in the linear maps, the distances between widely separated loci are not strictly proportional to the per cents of crossing over actually found, being relatively too large, in comparison with the per cents of crossing over. This he terms a "discrepancy" in the map, which has required the "subsidiary hypothesis" of double crossing over, in order to harmonize it with the theory of linear linkage. The answer to this is that it has never been claimed, in the theory of linear linkage, that the per cents of crossing over are actually proportional to the map distances: what has been stated is that the per cents of crossing over are *calculable* from the map distances—or, to put the matter in more mathematical terms, that the per cents of crossing over are a *function* of the distances of points from each other along a straight line. As will be shown presently, this circumstance alone is sufficient to

show that the factors must be bound together in a linear series; the precise nature of the function (involving coincidence, etc.) will then determine for us precisely the mode of incidence of the crossing over—*i. e.*, granted the linear series, it is then possible to calculate from the data the exact frequency of single crossing over, double crossing over of the various possible types, and multiple crossing over. Double crossing over thus becomes, not a "subsidiary hypothesis," but a phenomenon directly demonstrated.

It may, however, be noted in passing that, even if there had been no experimental evidence at all in regard to the nature of the linkage it could not have been conceded that Castle's alternative postulate—that no double crossing over can ever occur at all—would have been any more plausible *a priori* than that of the *Drosophila* workers which admits the existence of double crossing over. For, once the occurrence of single breaks in a chromosome is admitted—a point agreed upon by both sides—it is just as arbitrary to deny the possibility of double breaks as to assert their existence. Although Castle nowhere does explicitly admit that he has adopted this alternative "subsidiary hypothesis"—the denial of the possibility of double crossing over—yet an inspection of the theory of linkage which he himself has proposed shows that in this it has been tacitly assumed throughout, being necessary for the purposes of the solid models. Were double crossing over once admitted to occur, it could no longer be claimed that the distances between the factors in the three-dimensional models are exactly proportional to their per cents of separation—a condition which it is the sole aim of the existence of the models to fulfill.

We may now return to examine more carefully the main argument upon which linear arrangement and its corollaries (double crossing over, coincidence, etc.) is based. The fact previously stated that the linkage relations between the genes are such that they are all calculable from the positions of points in a linear series may also be expressed as follows: given any three linked factors, A, B and C, if any two of the linkages between them are known—say, the linkage AB and BC—then the third linkage—AC—is determined (the most convenient practical method for calculating it is to make use of the "curve of coincidence" of the particular chromosome). This is, two of the linkage values may be taken as "independent variables" and the third is then "dependent" on them—in this sense we may say that B is linked directly to A and to C, but that A is only linked to C through the linkage of each of these factors with B. Since this is true of any combination of three-linked factors (ABC, BCD, CDE, ACD, etc.) it can be shown that the factors are all linked together in chain arrangement, any one factor being linked directly to only two others (those which we may regard as being on either side of it), its linkage with the rest being entirely dependent on these intermediary linkages.[1] This remains true as a discovered mathematical fact of the linkage relationships, shown first in experiments of Sturtevant's designed to investigate the problem, and this is what the writer has designated as "the law of linear linkage." Whether or not we regard the factors as lying in an actual material thread, it must on the basis of these findings be admitted that the forces holding them linked together—be they physical, "dynamic" or transcendental—are of such a nature that each factor is directly bound, in segregation, with only two others—in bipolar fashion—so that the whole group, dynamically considered, is a chain. This does not necessarily mean that the spatial relations of the factors accord with these dynamic relations, for it is conceivable *a priori* that factor A might be far off from B, in another part of the cell, or that both might be diffused throughout the cell, and that they might nevertheless attract each other, during the segregation division, by some sort of chemical or physical influence. In the discussion that follows, no implication as to the actual physical arrangement of the genes is intended when the terms "linear series," "distance," etc., are used; these will refer only to the relations existing between the points in the linear map, which may be regarded merely as a mathematical mode of representation of the data themselves. It will be shown, however, at the conclusion of this article, that when the various conditions which have to be fulfilled at segregation are taken into consideration, any other explanation for these peculiarly linear linkage findings than an arrangement of the genes in a spatial, physical line proves to be hazardously fanciful.

---

[1] *I. e.*, all the linkages (factorial ($n-1$) in number) between the n factors in a group, can be shown to be dependent on (functions of) only $n-1$ "primary" or "independent" linkages. To obtain the most perfect expression of this dependency it is necessary to chose as the $n-1$ independent values the two strongest linkages involving each of the n factors (what we should call linkages AB, BC, CD, DE, etc., as contrasted with AC, AD, AE, BD, etc.). On this system, the other linkages all become definitely determined, the secondary linkages being in each case a function involving the sum of certain of the primary linkages. If, however, the primary linkages are not chosen according to the above rule, so as to constitute a "chain formation," no formula can definitely express the relationships of the linkages, for the secondary linkages will then in some cases depend upon the sum, in other cases upon the difference between the linkages taken as primary.

In the case of the larger distances, in order to discover what function of the distance the per cent. of separation represents, it would be necessary to conduct extremely delicate determinations, involving very extensive data in experiments dealing with many points simultaneously. Nevertheless enough has been done to show that even for the larger distances the per cents of separation do depend on the distance in the linear map—being less than the distance by an amount which varies in a fairly regular manner according to the distance itself; hence it is known that the higher per cents of separation certainly involve some function of the linear distance.

In the case of the smaller distances, on the other hand, the function has been rather accurately ascertained; it is very close to the simplest one possible, that is, there is an almost exact proportionality here between the map distances and the per cents of separations. Just as distance AB plus distance BC on a line are equal to distance AC, so the corresponding small frequency of separation between A and B, plus the small frequency between B and C, are found to be almost exactly equal to the frequency of separation between A and C; for this reason if the factors A, B and C are represented as points in a straight linear map, the distances between any two of them will represent the corresponding separation frequencies in an almost proportionate manner. A few examples of this principle are shown in Table I; it has been confirmed in innumerable other crosses, with many different factors. Moreover, it is found that the smaller the distance involved, the more exact is the porportionality that obtains, the less being the relative discrepancy between the frequency AC as found by experiment, and the value AC obtained, as on the map, by the summation of values AB and BC. The relationship which exists between the small separation values is hence just the sort which Castle himself would demand, for a proof of linear linkage. But whereas Castle would require this relationship to hold for all values, small or large, it may be shown that its existence in the case of the small alone is all that would be necessary for a complete proof of the doctrine of linear linkage, even if the large values were no sort of function of the linear series. For, if we proceed according to Castle's own method, and construct a map to represent the relations of the small values just described, showing each of the frequencies by a proportionate distance on the map, we necessarily obtain a map each section of which is practically a straight line. In the case, for example, of the data for v, g, and f, shown in Table I, if we represent the separation frequencies by proportionate distances in space, we must place point v at 10.7 units from g, and g at 11.3 units from f; if these two conditions are both to hold, then the only possible way of bringing f to its distance of 21.8 units from v is to put the three points in a nearly straight line, as shown

TABLE I

BACK CROSS OF FEMALES HETEROZYGOUS FOR v, g AND f. (PERFORMED BY BRIDGES; REPORTED BY WEINSTEIN)

I. Non-separations	II. Separations of v from g and f	III. Separations of f from v and g	IV. Separations of g from v and f	V. Total Flies
2651	360	380	3	3394

Resultant Per Cent. of Separations Between v and g	Resultant Per Cent. of Separations Between g and f	Resultant Per Cent. of Separations Between v and f
$\frac{(II + IV)}{V}$	$\frac{(III + IV)}{V}$	$\frac{(II + III)}{V}$
10.7	11.3	21.8

in Fig. 1. Other results indicate that the line would be exactly straight if still smaller distances were studied. Enough data have been obtained in the case of chromosome I of *Drosophila* to determine in this way the "shape" of each part of the linkage group, and each part, by itself, is thus found to follow the rules for linear

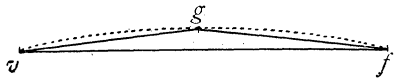

FIG. 1. Direct representation of the linkages in Table I. (*vg*, *gf*, and *vf* are each represented by a line of length proportionate to the respective frequency of separation.) The dotted curve shows the "average angular deviation" of the factors from a straight line.

distances in an extraordinarily rigorous manner. That is, given the factors ABCDE, etc.,—or to take an actual case, y, w, A, bi, cl,—it is found that the linkages of y, w, and A are proportional to their distances in a straight line, so are the linkages of w, A, and bi, for A, bi, and cl, etc. But, since every part of the group is thus linear, it must then be true that the entire group is linear. A line all of the parts of which are straight is a straight line. Any differences then observed between the size of the larger distances and the per cents of crossing over, even if they were so irregular that they could not be thought of as a function of the linear system itself, would then have to be regarded as due to peculiarities in the incidence of the crossing over, superimposed upon a system of genes which was really linear in formation,—

modifications due to specific correlations between crossings over in different regions. But since, as has been stated, the differences between the larger per cents of crossing over and the linear distances are not unregulated, but do give clear evidence of being themselves a function of the map distances, these larger per cents of separation as well as the smaller ones can be used in proof of the linear system of linkage. The systematic differences between the frequencies and the map are hence due to double and other multiple cross overs, which vary in frequency in accordance with the distance involved.

It is true that a certain amount of the differences actually found between the larger frequencies and their corresponding distances on a straight linear map might be thought of as due to the cumulation of minor discrepancies which existed between the small frequencies and distances but each one of which was by itself too small to be detected in the data for the smaller distances, being within the limits of experimental error. As the small discrepancies in such a case would, to be always cumulative, all have to have a bias in the same direction, this would amount to saying that the line along which the points were really disposed had a slow, even curve, too slight to be detected except when large distances were considered. The straight line would then be sufficiently accurate as a proportionate representation of all the

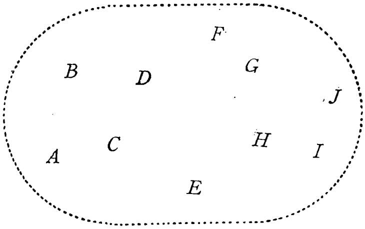

FIG. 2. Example of the sort of relations which would hold between factors arranged non-linearly (although in a plane figure). It is evident from this figure that more remote factors, such as *AE* and *J*, are likely to be arranged more nearly in a straight line than factors nearer together, such as *AB* and *C*.

small values but not of the large ones. If the validity of the evenly curved figure were accepted, it would in no way disagree with the finding of a linear arrangement of the genes, but would merely substitute a curved line for a straight one. In a really non-linear figure, such as shown in Fig. 2, the relations between the smaller distances would be (if anything) less of a linear type than the relations between the larger distances,—factors further apart in a thick rod, for example, would have to be more in line than those near together. The fact that the opposite relation holds in the actual data shows conclusively that the factors are in some sort of a line. There is an *a priori* objection, however, to accepting a curved line as an explanation of the linkage relations, in that it is very difficult to imagine a plausible set of conditions in the chromosome which would hold the factors rigidly in this curved line but which would at the same time determine the number of separations between the factors according to their direct (straight) distances from each other, instead of according to their distances along this line. But, quite aside from *a priori* reasons, there is an experimental result absolutely fatal to the curved line "explanation"; this consists in the finding of those classes which are termed by the *Drosophila* workers "triple crossovers." In the case of these classes the separations are of such a type as to require the assumption of a break in the curved line at three points simultaneously. As it is obvious that a break in only one plane could not cut the curve at more than two points, the triple crossovers therefore would have to be due to a break in more than one plane. The occurrence of breaks in more than one plane, however, disturbs the assumed relation of simple proportionality between separation frequencies and map distances, which was the basic postulate upon which the curved line was constructed. If the distances are after all once admitted to be not exactly proportional to the separation frequencies, then there remains no reason to assume, just because the larger separation frequencies are out of exact proportion to the distances on a straight linear map, that this is because the line is curved, and the factors thus nearer together. If still more evidence against the curved line idea be desired, it may be added that when the curve is constructed so as to be in good agreement (statistically) with the relations found between the smaller frequencies, it is then not sufficiently arched to permit the representation of the larger frequencies by proportionate distances (see section 4). A detailed compilation which I have made of all the data has shown that experimental error will not well account for the differences thus obtained between the two sets of results. The curved line being abandoned, it becomes, therefore, necessary to revert to "double crossing over," in explanation of the deviation of the large frequencies from the straight map values.[2]

3. If we examine further into Castle's argument, however, we find that he objects, not only because the larger separation frequencies are not proportional to the distances in linear maps, but also because he believes that the smaller frequencies are not proportional; in fact, according to his solid models, none of the kinds of frequencies, small or large, could even be a function of the distances in a linear map. In his solid, or rather, three-dimensional, models, which purport to have the factors so spaced that all distances between them are exactly proportional to the corresponding linkages, the factors are scattered about at all angles to each other, in such a

---

[2] The fact that the *geometrical* line which represents the linkages of the factors should be taken as straight does not imply that the supposed *physical* line in which the factors lie is straight. So long as the factors lie in any kind of physical line at all, then, if their linkages are determined, in some way, by their distances as measured *along* this line, these linkages should be representable on the basis of a straight geometrical map, inasmuch as all distances taken *along* a curved line must have the same interrelationships as distances in a straight line. Hence the curving of the chromosome filament is a matter entirely aside from the issue here involved, since the separation frequencies of the factors in the supposed filament are not conceived of as dependent upon their direct distances from each other but rather upon their distances along this filament. Thus the filament may, for these purposes, be treated as if it were straight.

way that their distances could never be represented as a function of distances in a single line. The cause of this discrepancy between Castle's figures and the relationships observed by the *Drosophila* workers lies in the nature of the data which Castle uses, or rather, in his manner of using the data. For Castle constructs his maps, or models, on all the data obtainable, indiscriminately, and regardless of the fact that most of the data for the linkage values involved have been secured in as many different experiments. On the contrary, it is necessary, in order to determine exactly the relationships existing between interdependent linkage values, that all the data be obtained from the same experiment. This is because the precise value obtained for any given linkage is not only subject to the ordinary error of random sampling but may vary significantly in different experiments, in response to different environmental conditions, the age of parents, genetic factors, and the amount of discrepancy due to differential viability. Piling up enormous counts does not eliminate these sources of variation. Any slight aberration thus produced in the absolute value of one of the linkages (say AC) will then alter so materially its relative value, as compared with the other linkages (AB and BC), obtained in two different experiments, that the different values no longer fit into the linear system; they will not be expressible as any sort of function of the system. That is why Castle found that the *Drosophila* workers' own data gave the per cent. of crossing over between y(yellow) and w(white) as 1.1, between w and bi (bifid) as 5.3, and between y and bi as 5.5, a relationship quite at odds with their claims concerning linear linkage for short distances. Castle could have pointed out numerous similar "discrepancies," by similarly choosing to compare exactly (within, say, one unit of distance), the results of different experiments. In fact, had we been allowed to select the experiments for him, we could have chosen values such as the following: Sb (frequency between star and black) 39.3;

bp (frequency between black and purple) 5.9; Sp 0.4. If Castle will follow his usual procedure here, and represent these frequencies by proportionate distances in a model, he will disprove not only linear linkage but both Euclidean and non-Euclidean geometry and plain arithmetic. The trouble in the case just cited arises in the fact that the first two values are those obtained under ordinary circumstances whereas the third is a value obtained in the presence of the factor CIIL which decreases enormously the amount of crossing over. Clearly it will be unfair to expect a single map to represent all three values simultaneously. Nevertheless, similar although less exaggerated, disturbing influences may be, and frequently are, at work causing discrepancies between the results of "ordinary" experiments, so that it should be evident that the latter are not ordinarily fit to be subjected to the delicate comparison which is necessary for the purpose of determining the nature of the linkage system.

To some critics, it might at first sight appear inconsistent for the *Drosophila* workers to use the above argument against Castle's system, in view of the fact that these workers themselves also combine the results of different experiments in constructing their chromosome maps. The answer to this is that the variations in linkage between ordinary experiments are usually so small absolutely, that, if all the data for independent linkage values, like AB, BC, CD, etc.—are joined together and represented in one linear map, the latter will be accurate enough for the usual purpose of computing approximately the per cents of separation: the factors will appear in their correct order, and with approximately the correct distances between them. If, however, a study of the nature of the system of linkage is to be made, much more precise knowledge than this is required, for it is necessary to know exactly the relative strengths of interdependent linkages—like AB, BC and AC—as compared with one another. In such a case the small absolute deviations occurring in the different experiments become large relative deviations of the linkages as compared with each other—this is particularly true the smaller the absolute per cents of separation are—and so a totally erroneous impression of the nature of the linkage system may be produced. The nature of the linkage system—whether it is linear and, if so, what function of a line is involved—can only be studied to the best advantage in experiments involving several factors at the same time, but if our judgments regarding it have already been arrived at, or verified, in this way, it is then quite legitimate to use this knowledge for other factors, and to join the results of different experiments involving them into one linear map.

TABLE II

Separation Frequencies Between Every Two of the Six Sex-linked Factors y, bi, cl, v, s, B, as shown in a Count of 712 Flies from a Cross in which Twelve Sex-linked Factors were Followed Simultaneously. (Muller)

Directness of the Linkage, According to the Linear Map	Factors Considered	Number of Separations	Per Cent. of Separations	Sum of Numbers of Separations of Each from an Intermediate Factor	Sum of Per Cents. of Separations of Each from an Intermediate Factor
"Primary" (for the experiment)	y and bi	39	5.5		
	bi and cl	53	7.4		
	cl and v	112	15.7		
	v and s	57	8.0		
	s and B	95	13.3		
Dependent on two "primaries"	y and cl	92	12.9	y bi + bi cl = 92	12.9
	bi and v	165	23.1	bi cl + cl v = 165	23.1
	cl and s	167	23.4	cl v + v s = 169	23.7
	v and B	152	21.3	v s + s B = 152	21.3
Dependent on three "primaries"	y and v	198	27.8	y cl + cl v = 204	28.6
	bi and s	216	30.3	bi v + v s = 222	31.2
	cl and B	240	33.7	cl v + v B = 264	37.0
Dependent on four "primaries"	y and s	247	34.7	y cl + cl s = 259	36.4
	bi and B	275	38.6	bi v + v B = 317	44.5
Dependent on five "primaries"	y and B	296	41.6	y v + v B = 350	49.2

In the experiments previously cited, the nature of the linkage in various sections of the chromosome has been studied by following the inheritance of three factors in that region simultaneously. By a series of extensive counts of this sort the nature of the linkage in each individual section of the first chromosome has been studied, and found to be linear. The data in Table II are derived from an experiment which involves less extensive numbers than these, but illustrates to better advantage the linear behavior of all parts of the chromosome at once. These data are taken from Muller's cross of flies heterozygous for twelve mutant sex-linked factors. The results for six of these factors—those scattered most evenly along the chromosome—are shown in the table, which gives the number and per cent. of separation between every one of these factors and each of the other five. It will be seen that it happened that in this particular experiment, for all per cents of separation below 23, the per cent. of separation between any two factors was exactly equal to the sum of the per cents of separation of each from a third factor lying between them, whereas for factors less closely linked, the larger per cent. was less than the sum of the other two by an amount varying closely with the size of the large frequency itself. The data obtained in this same experiment for the three factors y, w and bi are given separately in Table III, in order that they may be compared to better advantage with the non-linear relation for these factors which Castle claims, as a result of his combination into one map of the results of separate experiments. Whereas Castle obtained a triangular figure to represent the three frequencies (y w 1.1, w bi 5.3, and y bi 5.5) it is seen that in this experiment, where all three were followed at the same time, an exactly linear relationship was obtained (y w 1.7, w bi 3.8, y bi 5.5). An experiment of Sturtevant's involving just these three factors is shown in the same table (III); here too the relations are entirely linear. In like manner the values obtained in the 12-factor experiment for the loci of y w and A are given in Table IV (y w 1.7, w A 1.4, y A 3.1) to be compared by the "triangular" values (y w 1.1, w A 1.7, y A 2.0) claimed by Castle. The numbers in these experiments are quite sufficient to have revealed clearly any such triangular relationships as shown in the data chosen and figured by Castle.

TABLE III

SEPARATION FREQUENCIES OF y, W AND bi

A. Data from the same experiment as that which furnished Table II. (Muller)

I. Non-separations	II. Separations of y from w and bi	III. Separations of bi from y and w	IV. Separations of w from y and bi	V. Total Flies
673	12	27	0	712

Resultant Per Cent. of Separations Between y and w	Resultant Per Cent. of Separations Between w and bi	Resultant Per Cent. of Separations Between y and bi
$\frac{(II + IV)}{V}$	$\frac{(III + IV)}{V}$	$\frac{(II + III)}{V}$
1.7	3.8	5.5

B. Data from a cross involving just these three factors. (Sturtevant)

I. Non-separations	II. Separations of y from w and bi	III. Separations of bi from y and w	IV. Separations of w from y and bi	V. Total Flies
487	3	16	0	506

Resultant Per Cent. of Separations Between y and w	Resultant Per Cent. of Separations Between w and bi	Resultant Per Cent. of Separations Between y and bi
$\frac{(II + IV)}{V}$	$\frac{(III + IV)}{V}$	$\frac{(II + III)}{V}$
0.6	3.2	3.8

TABLE IV

SEPARATION FREQUENCIES OF y, W AND A.

(From the same experiment as that which furnished Tables II and IIIA. Muller)

I. Non-separations	II. Separations of y from w and A	III. Separations of A from y and w	IV. Separations of w from y and A	V. Total Flies
690	12	10	0	712

Resultant Per Cent. of Separations Between y and w	Resultant Per Cent. of Separations Between w and A	Resultant Per Cent. of Separations Between y and A
$\frac{(II + IV)}{V}$	$\frac{(III + IV)}{V}$	$\frac{(II + III)}{V}$
1.7	1.4	3.1

4. Although it has been shown that the linkage relations existing among the factors in any one experiment are functions of a linear series it might still be questioned whether there might not, after all, be some advantage in using Castle's system of graphic representa-

tion—whereby each separation frequency is supposed to be shown by an exactly proportionate distance on the figure, no matter how many dimensions may be required for this purpose. It will now be shown, however, that such a system of representation is impossible, quite aside from the fact that the models shown in Castle's papers are based upon data which can not legitimately be combined together. That is, no matter whether the data used are all derived from one experiment, or whether the results of different experiments are combined according to Castle's method, they could not be represented either in a three-dimensional or in any other geometrical figure, in such a way that all the distances would be proportional to the separation frequencies.

This may be illustrated by the data reported in Table II. It has been seen that the per cent. of separations between y and cl is exactly equal to the per cent. of separations between y and bi plus that between bi and cl. If then we represent these frequencies by actual distances, we must make the distance between points y and cl exactly equal to the distance between y and bi plus that between bi and cl. The only possible way to do this, on any kind of geometry—one-dimensional, three-dimensional or *n*-dimensional—it to put these three points in one straight line. In a similar manner we must place bi cl v in a straight line, and also v s B. Cl v and s are in almost a straight line, but there would have to be a slight bend at v, owing to the fact that cl s is very slightly shorter than cl v plus v s (on account of just one double crossover having occurred between them); this is correlated with the fact that cl s is a longer distance than the others considered. The figure so constructed, on the basis of Castle's own methods, is shown in Fig. 3; it is quite evident that this is the only figure which will represent directly (proportionately) the frequencies above considered. If, however, we now measure the distance on this figure between the extreme points, y and B, we find that it turns out to be 49.3, or very nearly the sum of the intermediate distances (50.0), whereas the frequency of separation found between y and B in the actual experiment is 41.6. Similarly, the "model" shows too high a frequency for the other longer distances involved. (The long distances y s and bi B are 36.2 and 43.9 respectively on the model, but only 34.6 and 38.6 in the data; the moderately long distances y v, bi s, and cl B are 28.6, 30.8, and 36.6, respectively, on the model, but 27.8, 30.3, and 33.7 in the data.) It would, on the other hand, have been possible to bring y and B close enough together in the diagram, and at the same time have adjacent factors the correct distance apart, by giving the line a curve, or bending it, as shown in Fig. 4. But if this is done it is found that the distances of intermediate length (y cl, bi v, cl s, and v B) are not properly represented, all of them being relatively too short on the diagram. It would be unsafe to attribute these discrepancies, so uniform in direction, to the "errors of random sampling." The present experiment is cited, however, purely

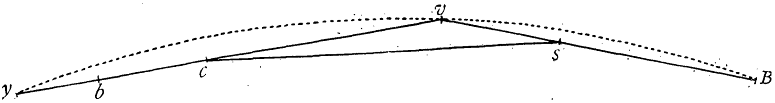

FIG. 3. Direct representation of the strongest and second strongest linkages in Table II. (*y bi, bi cl, cl v, v s, s B,* and *y cl, bi v, cb s, v B,* are each represented by a line of length proportionate to the respective frequency of separation.) The dotted curve shows the "average angular deviation" of the line of factors, according to this system.

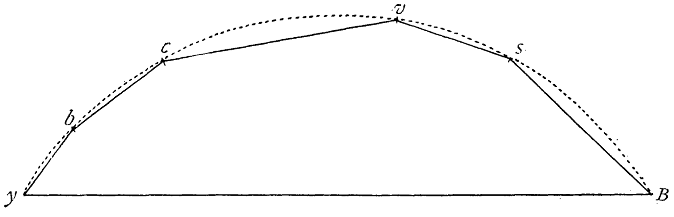

FIG. 4. Direct representation of the strongest and weakest linkages in Table II. (*y bi, bi cb, cb v, v s, s B,* and *y B* are each represented by a line of length proportionate to the respective frequency of separation.) The dotted curve shows the "average angular deviation" of the line of factors, according to this system.

as an illustration, to show what kind of discrepancies are meant. Discrepancies of exactly the same character and direction appear when the diagrams obtained by this method from experiments involving three points close together are compared with those from other experiments having three points far apart; that is, the former figures are repeatedly found to be nearer in form to a straight line than the latter; in such three-point experiments, moreover, highly extensive counts have been made, involving altogether (in the published experiments on the first chromosome), approximately a hundred thousand flies, and thousands of double cross overs.

5. The relation which has just been described, whereby the larger frequencies of separation are relatively smaller than could be directly represented in a curve constructed on the basis of the small frequencies, is due, according to the phraseology of the *Drosophila* workers, to the fact that the relative frequency of double crossing over ("coincidence") is so much larger for large frequencies than for small ones. Castle realizes to a certain extent the difficulty which this circumstance entails for his models, and he endeavors to meet it by means of the "subsidiary hypothesis" that the breaks in his models are more frequent in certain directions than in others. This assumption would, in some measure, explain away in a formal manner certain of the discrepancies (although cases of "triple crossing over" still remain an insurmountable obstacle), but the adopting of any such hypothesis really amounts to cutting away the ground from under the main theory of "proportionate representation," for the hypothesis involves an abandonment of the claim that the model represents each frequency by a proportionate distance between the nodes. For it is evident that if, in a given region, breaks in one direction are more frequent than in another, then points in this region which are an equal distance apart will be separated with different frequency according to the di-

rection of the line joining them. A given distance then no longer represents a given frequency.

6. It has been shown in the above two sections that a single figure will not represent accurately, by proportionate distances, the various linkage frequencies actually found in experiments involving many factors at once. If, on the other hand, it had been attempted to combine into one map the absolute frequencies obtained in a series of different experiments with two factors at a time, as Castle claims to do, the number and extent of discrepancies irreconcilable with any possible geometrical figure would have been much greater still. For, since the absolute frequencies found in different experiments necessarily have all sorts of irregular relationships to each other, it follows that it would be even less possible to show in one solid model separation frequencies which were obtained in this way. One such irreconcilable value has already been recognized by Castle—namely, the frequency y B. He is forced to represent this frequency by a curved wire in his model, because it is longer (being 47) than the longest distance possible (41) between these two points in any figure founded on a proportionate representation of the other frequencies. Of course, if distance on the model is to have any meaning, it cannot arbitrarily be represented along a straight line in some instances, and along lines having various degrees of curvature in other instances. In this one case, then, Castle is compelled to assume that the wrong value has been obtained, owing to experimental error, even though in all other cases he has assumed that it is quite legitimate to combine the results of different experiments. It would have been strange if, in making a model of this sort representing the separation frequencies of so many combinations of factors, Castle had not encountered more of these refractory cases. He does not mention any more, but it is noticeable that several other curved lines appear in his model. Moreover there is a conspicuous absence in the model of the factor lethal 2.

It would have been inconvenient to represent the observed linkage results for this factor by proportionate lines in the model, for, according to the results, lethal 2 would have to be placed at 9.6 from w and 17.7 from v, thus making the distance between w and v not greater than 27.3, whereas the established distance between w and v themselves is as much as 30.5; similarly, although only 9.6 from w, lethal 2 must be placed only 15.5 from m, although w and m are known to be at least 33.2 units from each other. Either the wv and wm lines would have to be considerably curved, therefore, or the lines between lethal 2 and the other factors would have to be *stretched* in some way—perhaps dotted lines would meet the difficulty!

Although any scheme of representing linkage results by exactly proportionate distances encounters the contradictions discussed above, it is noticeable that nearly all the most extreme departures from a plane curved figure (that figure which *comes nearest* to representing, by strictly *proportionate* distances, the ratios resulting from that type of linear linkage which actually exists) occur in the case of factors whose linkage ratios are distorted by differential viability or difficult classification. This is why the factors A, fr, sh, cl, bi, and the lethals stand out from the fairly regular curved line which Castle's models would otherwise conform to. The first two of the above factors are uncertain of classification, the others mentioned affect viability markedly. In the case of nearly all the remaining factors of the model, even though the results were taken from different experiments, it was nevertheless found, when the data were plotted, that they agreed pretty well with the expectations based on linear linkage. Moreover, the better the experimental conditions are in regard to viability, certainty of classification, and size of counts,—the more closely are these synthesized results of individual experiments found to coincide with the linear findings of experiments involving three or more points simultaneously. (It is for this reason that Sturtevant was first able to hit upon the general fact of linear linkage, on the basis of numerous careful experiments involving only two factors at a time.)

7. Owing to the inherent inconsistencies of the methods that were used to construct the solid models, it is to be expected that any predictions regarding separation frequencies which are deduced from them would be extremely unsafe. Castle states, however, that if any newly discovered gene has been located in the model, by obtaining its frequencies of separation from any three of the other genes contained therein, then the relation of the new gene "to all the others could be predicted by direct measurement from the model." In the case of two of the four predictions which Castle has made in this way, some evidence concerning the distance between the loci involved is already in existence.

One of the frequencies of separation in question is that between the loci of the recessive mutant factors glazed eye and rugose eye, in *Drosophila virilis*. Castle predicts, on the basis of his model, that the per cent. of crossing over between them should be found to be 4 or 5, or "probably a little greater." The work of Metz, and unpublished work of Weinstein, have shown, however, that hybrid females which carry both mutant factors exhibit the somatic character sterility possessed by the more extreme mutant type. When this dominance of an ordinarily recessive character in F_1 is taken together with the close similarity between the unusual effects produced by the two mutant factors (both produce a similar, peculiar effect on the eye, which is sex-limited, being more marked in the males), and with the fact that there is also a third mutant member of the series, with similar peculiar effects and similar linkage relations, it becomes highly probable that these factors are all allelomorphs. In that case they occupy "identical loci," and the frequency of separation between them must be 0. A direct determination of the per cent. of crossing over between them is obviously impossible to obtain, on account of the sterility of the females carrying both factors.

Another prediction based on the solid models dealt with the frequency of separation between the factors hairy and magenta, of *Drosophila virilis*. It had been found by Metz that the per cent. of cross overs between hairy and forked was 3.1 and between forked and magenta 3.7; this would make the per cent. for hairy magenta 6.8 (or 0.6), if the factors were in a straight line. In the solid model, however, the arrangement of these factors, based on separate determinations of their frequencies of crossing over with distant loci,—is shown as triangular; and on the basis of this model Castle predicts that the frequency hairy-magenta will be found to be 4 or 5. The frequency has recently been determined by Weinstein, who has kindly consented to allow its use in this connection. He finds it to be 6.6.

It should be pointed out that Castle has endeavored to protect himself, in these predictions, by saying that they only hold, provided the relations given "have been determined with sufficient accuracy"!

8. One of Castle's specific objections to the linear maps, on which he lays much stress, is that on them the distances between the extreme factors is much more than 50 units, whereas factors which are linked must have a separation frequency of less than 50 per cent. It is only necessary to point out here that since, as we have seen, the linear maps, unlike the models, do not imply a proportionate relationship between the distances and the separation frequencies, these distances of over 50 do not connote separation frequencies of over 50 per cent. On account of the progressive reduction in separation frequency, due to double crossing over, that occurs with increasing distance, even distances of 100 or 110 in the second chromosome do not connote separation frequencies as high as 50 per cent. On the other hand, it should also be remarked that separation frequencies of over 50 per cent. would not be impossible *a priori*, as Castle

maintains; consequently any system of representing linkage which permitted or showed such values would not be *ipso facto* inconsistent. The mere fact that all factors hitherto worked with in a single chromosome have less than 50 per cent. of separation, and that those in different chromosomes have just 50 per cent., does not mean that factors can never be found which are so far apart, and which lie in such a rigid chromosome (little double crossing over) that they separate more often than they remain together at segregation. Whether this phenomenon should then be called linkage is but a question of words; the chromosomes themselves would have no regard for the 50.0 per cent. mark, or for the idiosyncrasies of our terminology.

The proof of the law of linear linkage, including all the main aspects of it which have been given above, has been stated on several previous occasions. It seems unfortunate that the argument has had to be repeated each time that a new "theory of crossing over" has arisen, for the discussion and data given in the original papers supply all the material necessary for a decision of the matter, at least so far as the germ plasm of *Drosophila* is concerned.

Before closing, it may be desirable to supplement these arguments for a mathematically linear mode of linkage, by a statement of the considerations which indicate that this mathematically linear linkage can have its basis only in a linear physical connection between the genes.

If the genes are not spatially arranged, or physically connected, in the same linear sequence as that in which they have been found to attract each other in linkage, then the forces of linkage attraction must be such as to "act at a distance." But, although acting at a distance, these linkage forces must nevertheless be extraordinarily specific—binding each gene directly to just two specific associate genes. Hence the forces could not be of an electrical nature, for, since there are only two kinds of electricity, electric forces could not be specific enough.

Similarly, the attractions could not be magnetic, nor could they be due to any kind of diffuse "physical" forces, such as those that emanate from centers of surface tension change or from centers of vibrational disturbances. Those who deny linear arrangement, while admitting the mathematically linear linkage results would therefore be driven to assume that the linkage attraction depended on the specific chemical nature of the genes, which, by virtue of their chemical composition, exerted a specific attraction at a distance, as the substances of adsorption compounds are sometimes supposed to do. But such a theory, as a method of accounting for linkage, becomes stretched to the breaking point when it is remembered that each gene must be assumed to have such an attraction for just two of the others, never more nor less, and that when this attraction is broken it is always exchanged for that of the allelomorph. Moreover, it would be exceedingly hard to reconcile this theory with the finding that changes in the nature of the genes—mutations—alter in nowise the sequence of their linkage attractions, and very rarely change even the strength of the linkages. And when we come to analyze the linkage relations in detail, and encounter the phenomenon of interference, we find relations that are entirely at variance with all our preconceptions concerning chemical attractions or chemical activity in general,—results that would force us to assume (1) that a breakage of the attraction between two genes leads to an increased attraction between the other genes and (2) that the amount of this increased attraction ("interference") depends solely on the directness of the connection ("distance") between these other genes and the one whose attraction was broken, being not at all influenced by the chemical nature of the broken attraction, or by the chemical nature of the other attractions themselves. The facts of "interference" or "coincidence" are thus diametrically opposed to a chemical view of linkage, although they, like all the other facts of linkage, are quite in accord with ideas of a spatial, physical linear arrangement, their interpretation on the latter basis being natural and obvious.

The idea that the genes are bound together in line, in order of their linkage, by material, solid connections thus remains as the only interpretation which fits the genetic findings. In view of the additional fact that the chromosomes—themselves known to be specifically linked to the factor groups—can, at certain stages of their history, be seen to have the linear structure required, it would indeed be rash to adopt a different theory, without most cogent evidence of a startlingly new character.

BIBLIOGRAPHY

Castle, W. E.

1919. Is the Arrangement of the Genes in the Chromosome Linear? *Proc. Nat. Acad. Sc.*, V, 25–32.

1919. The Linkage System of Eight Sex-linked Characters of *Drosophila virilis*. *Proc. Nat. Acad. Sc.*, V, 32–36.

Metz, C. W., and Bridges, C. B.

1917. Incompatibility of Mutant Races in *Drosophila*. *Proc. Nat. Acad. Sc.*, III, 673–678.

Metz, C. W.

1918. The Linkage of Eight Sex-linked Characters in *Drosophila virilis*. *Genetics*, III, 107–134.

Morgan, T. H., and Bridges, C. B.

1916. Sex-linked Inheritance in *Drosophila*. Carnegie Institution of Washington Publ. 237, 88 pp.

Morgan, T. H., Sturtevant, A. H., Muller, H. J., and Bridges, C. B.

1915. The Mechanism of Mendelian Heredity. 262 pp. Henry Holt.

Muller, H. J.

1916. The Mechanism of Crossing Over. AMER. NAT., I, 193–221, 284–305, 350–366, 421–434.

Sturtevant, A. H.

1915. The Behavior of the Chromosomes as Studied through Linkage. *Zeitschr. f. ind. Abst. u. Vererb.*, XIII, 234–287.

Weinstein, A.

1918. Coincidence of Crossing Over in *Drosophile melanogaster* (*ampelophila*). *Genetics*, III, 135–159.